Assignment - 1

2022-09-10

Import Dataset

```
options(stringsAsFactors = FALSE)
Sales_Data <- read.csv("~/Downloads/Avinash/Avinash/Sales_Data.csv")
head(Sales_Data)
```

```
##   Sales_Rep Business Age Female Years College Personality Certficates Feedback
## 1         1 Hardware  59      1     2     Yes    Diplomat           1     2.01
## 2         2 Hardware  52      0    10     Yes    Diplomat           4     3.64
## 3         3 Software  47      1     1     Yes    Explorer           1     3.88
## 4         4 Hardware  61      0     2     Yes    Diplomat           3     2.70
## 5         5 Software  39      0     1      No    Diplomat           5     3.44
## 6         6 Hardware  28      0     6     Yes    Explorer           1     2.43
##   Salary NPS
## 1  70200   5
## 2 133000  10
## 3  52600   8
## 4  96000   6
## 5 122000   7
## 6  60000   6
```

Descriptive statistics

```
library(pastecs)
stat.desc(Sales_Data[,c("Sales_Rep","Business","Age","Certficates","Salary")])
```

```
##                 Sales_Rep Business          Age  Certficates       Salary
## nbr.val      2.199000e+04       NA 2.199000e+04 2.199000e+04 2.199000e+04
## nbr.null     0.000000e+00       NA 0.000000e+00 2.113000e+03 0.000000e+00
## nbr.na       0.000000e+00       NA 0.000000e+00 0.000000e+00 0.000000e+00
## min          1.000000e+00       NA 2.100000e+01 0.000000e+00 2.100000e+04
## max          2.199000e+04       NA 6.500000e+01 6.000000e+00 1.970000e+05
## range        2.198900e+04       NA 4.400000e+01 6.000000e+00 1.760000e+05
## sum          2.417910e+08       NA 9.124960e+05 5.744200e+04 1.620086e+09
## median       1.099550e+04       NA 4.100000e+01 2.000000e+00 7.000000e+04
## mean         1.099550e+04       NA 4.149595e+01 2.612187e+00 7.367378e+04
## SE.mean      4.280868e+01       NA 7.696897e-02 1.111508e-02 1.535496e+02
## CI.mean.0.95 8.390810e+01       NA 1.508647e-01 2.178636e-02 3.009682e+02
## var          4.029851e+07       NA 1.302736e+02 2.716755e+00 5.184686e+08
## std.dev      6.348111e+03       NA 1.141375e+01 1.648258e+00 2.276991e+04
## coef.var     5.773371e-01       NA 2.750569e-01 6.309878e-01 3.090639e-01
```

Transformation

```
library(tidyverse, quietly=TRUE, warn.conflicts=FALSE)
```

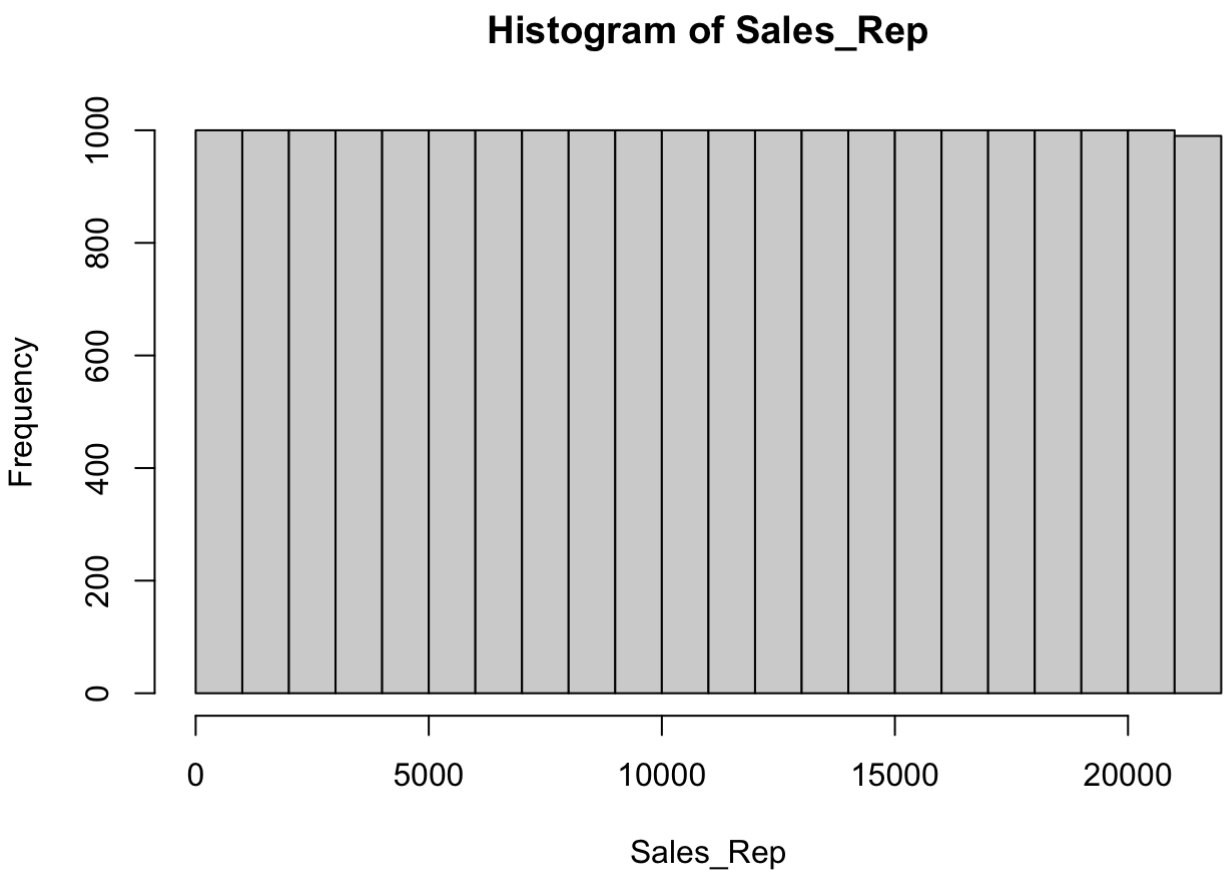
```
## ── Attaching packages ─────────────────────────────────────── tidyverse 1.3.2 ──
## ✔ ggplot2 3.3.6      ✔ purrr   0.3.4
## ✔ tibble  3.1.8      ✔ dplyr   1.0.10
## ✔ tidyr   1.2.0      ✔ stringr 1.4.1
## ✔ readr   2.1.2      ✔ forcats 0.5.2
## ── Conflicts ────────────────────────────────────────── tidyverse_conflicts() ──
## ✖ tidyr::extract() masks pastecs::extract()
## ✖ dplyr::filter()  masks stats::filter()
## ✖ dplyr::first()   masks pastecs::first()
## ✖ dplyr::lag()     masks stats::lag()
## ✖ dplyr::last()    masks pastecs::last()
```

```
Arrange_Age <- Sales_Data %>% arrange(Age)
head(Arrange_Age)
```

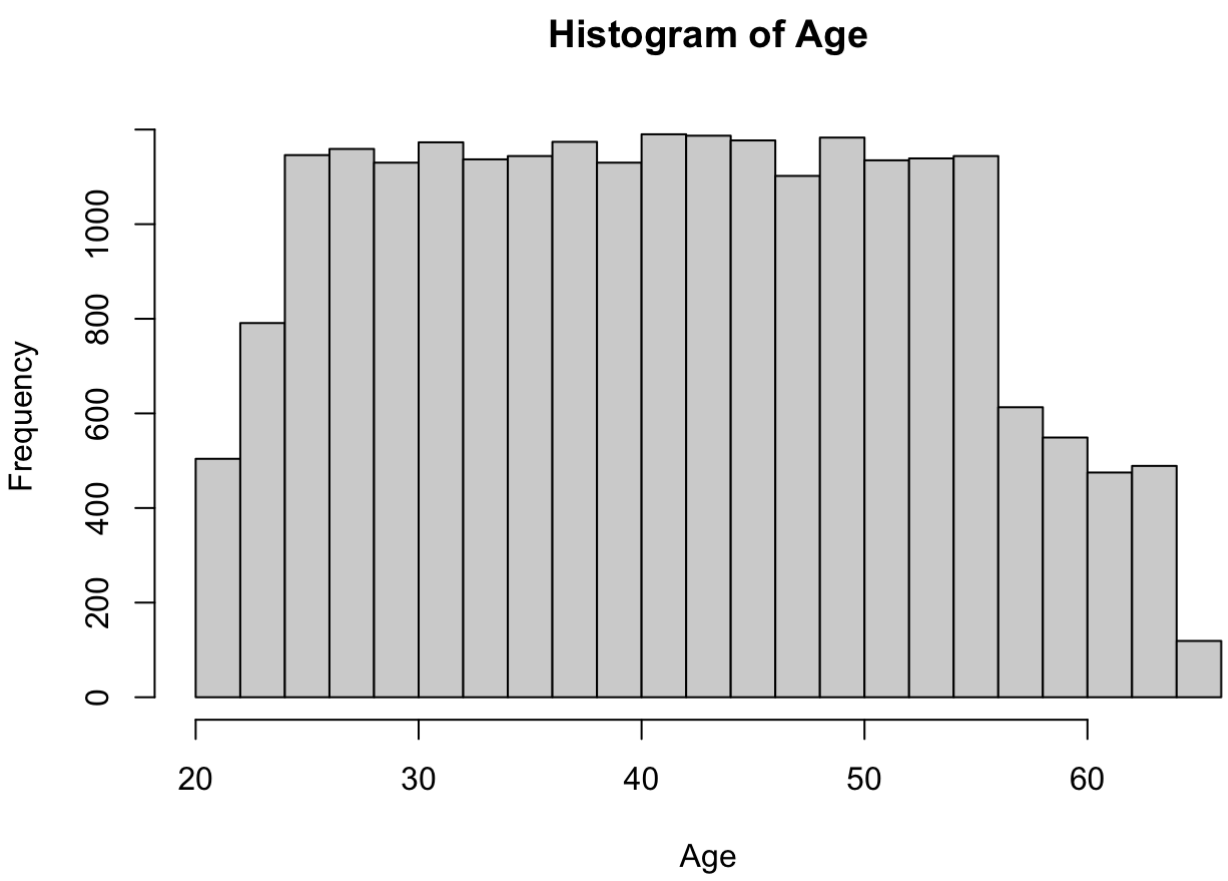
```
##   Sales_Rep Business Age Female Years College Personality Certficates Feedback
## 1       157 Software  21      0     2     Yes     Analyst           4     1.52
## 2       655 Software  21      1     1     Yes    Explorer           3     2.04
## 3       735 Software  21      0     2     Yes    Sentinel           2     3.51
## 4       752 Software  21      1     4      No     Analyst           3     2.90
## 5       849 Software  21      0     1     Yes    Sentinel           0     2.36
## 6       925 Software  21      1     3     Yes    Diplomat           3     3.08
##   Salary NPS
## 1  46000   3
## 2  71800   4
## 3  50000   7
## 4  41800   4
## 5  27000   3
## 6  55800   7
```

Quantitative plots

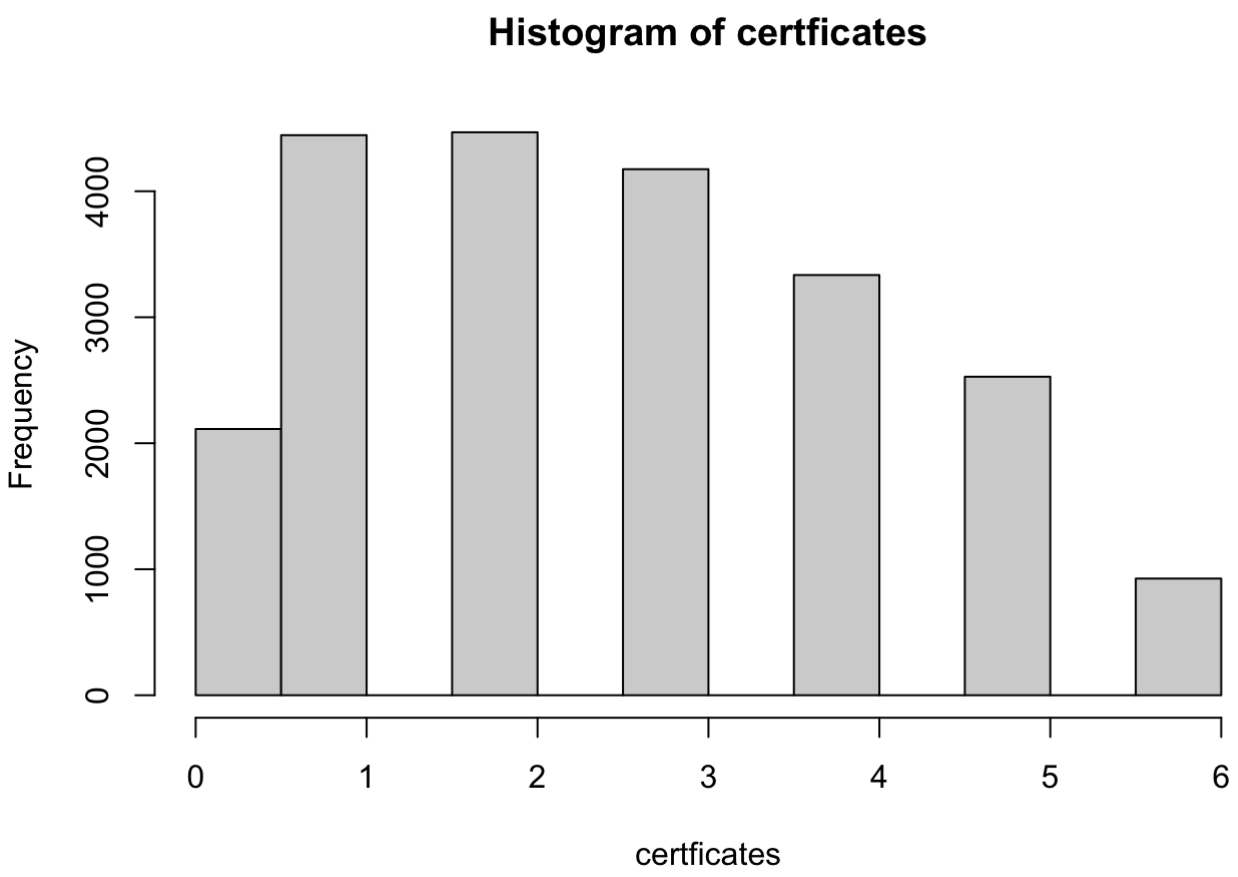
```
Sales_Rep <- Sales_Data$Sales_Rep
Business <- Sales_Data$Business
Age <- Sales_Data$Age
certficates <- Sales_Data$Certficates
Salary <- Sales_Data$Salary
hist(Sales_Rep)
```


Histogram of Sales_Rep
Frequency
0
200
400
600
800
1000
0
5000
10000
15000
20000
Sales_Rep

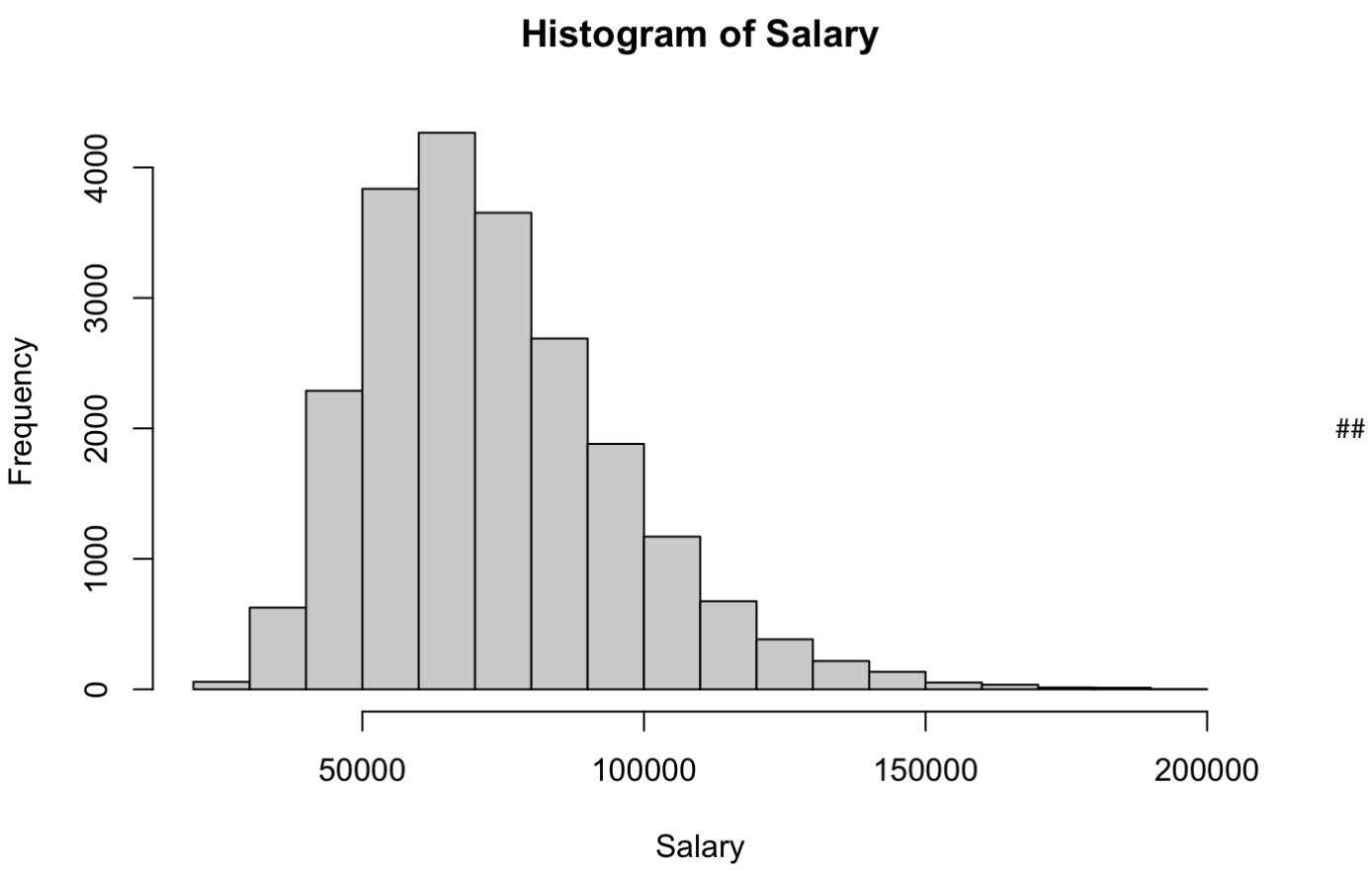
```
hist(Age)
```




```
hist(certficates)
```


Histogram of certficates
Frequency
0
1000
2000
3000
4000
0
1
2
3
4
5
6
certficates

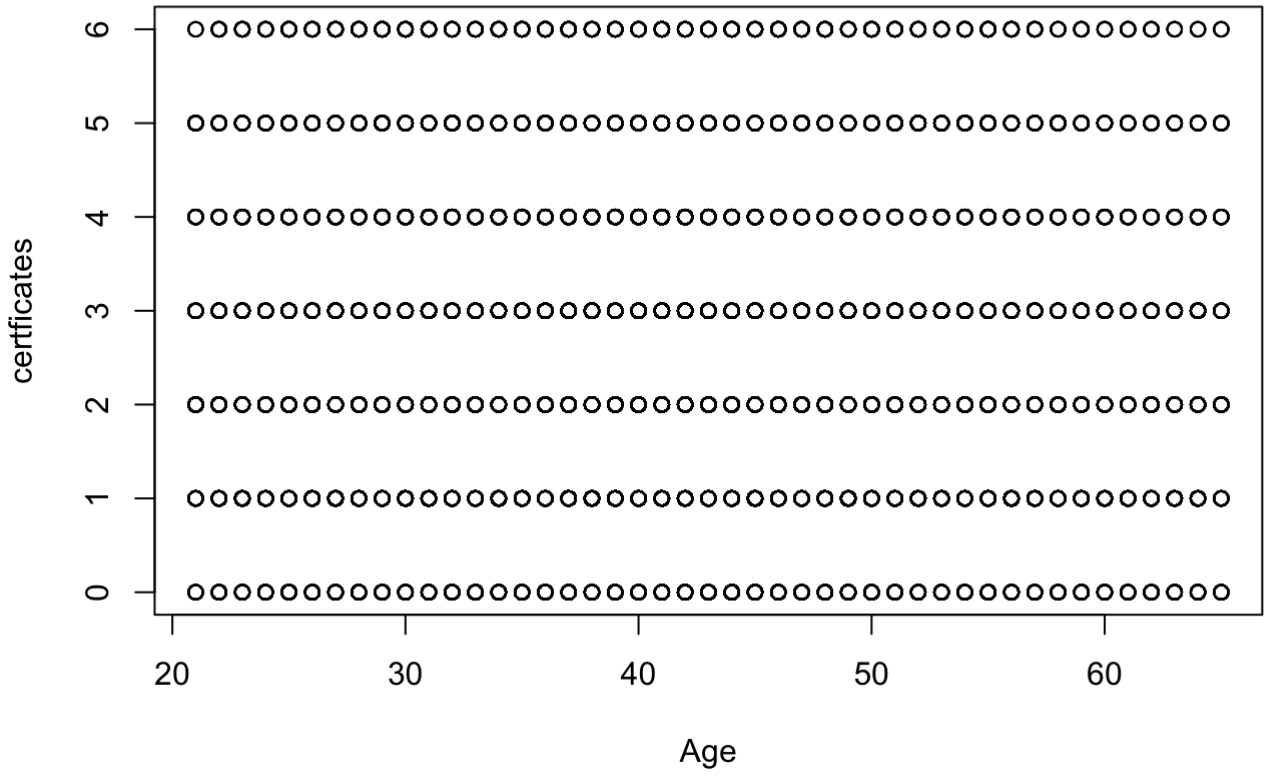
```
hist(Salary)
```



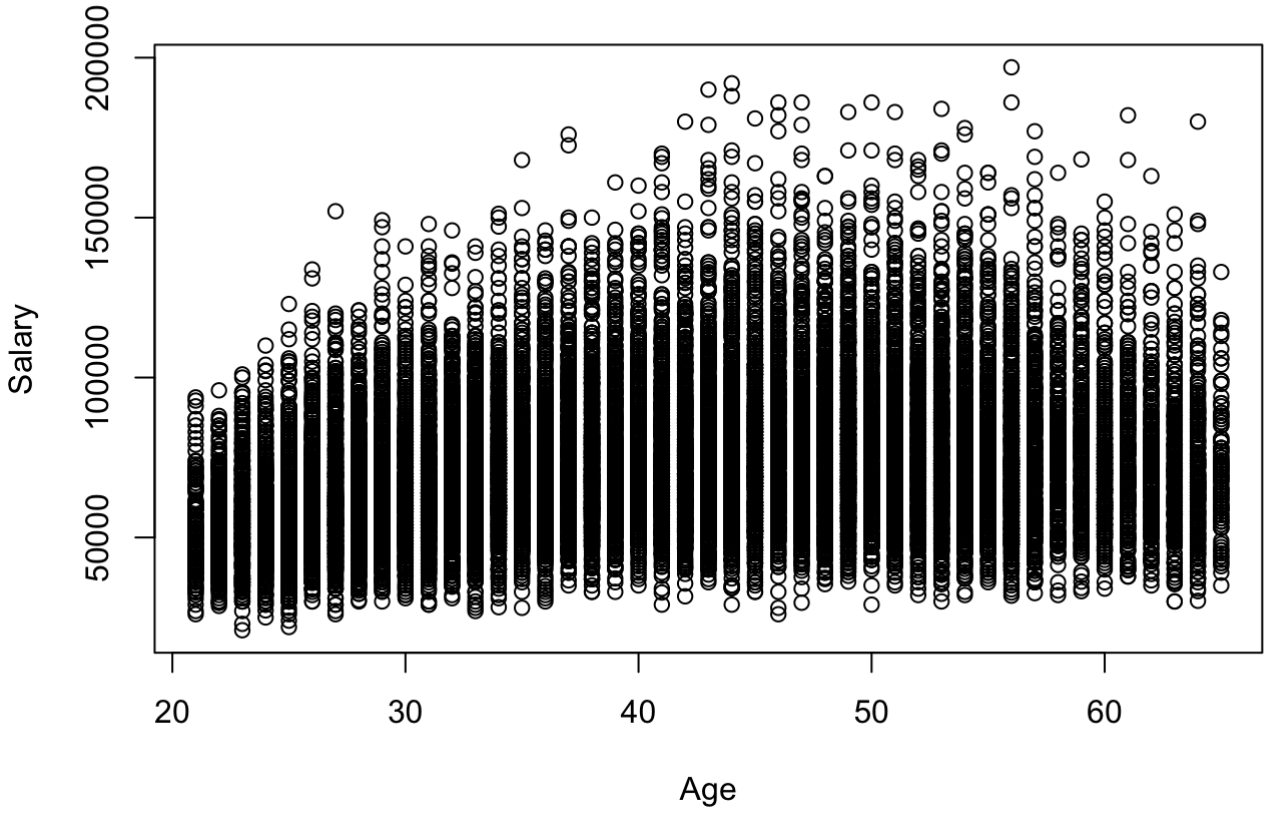
##

Scatter plots

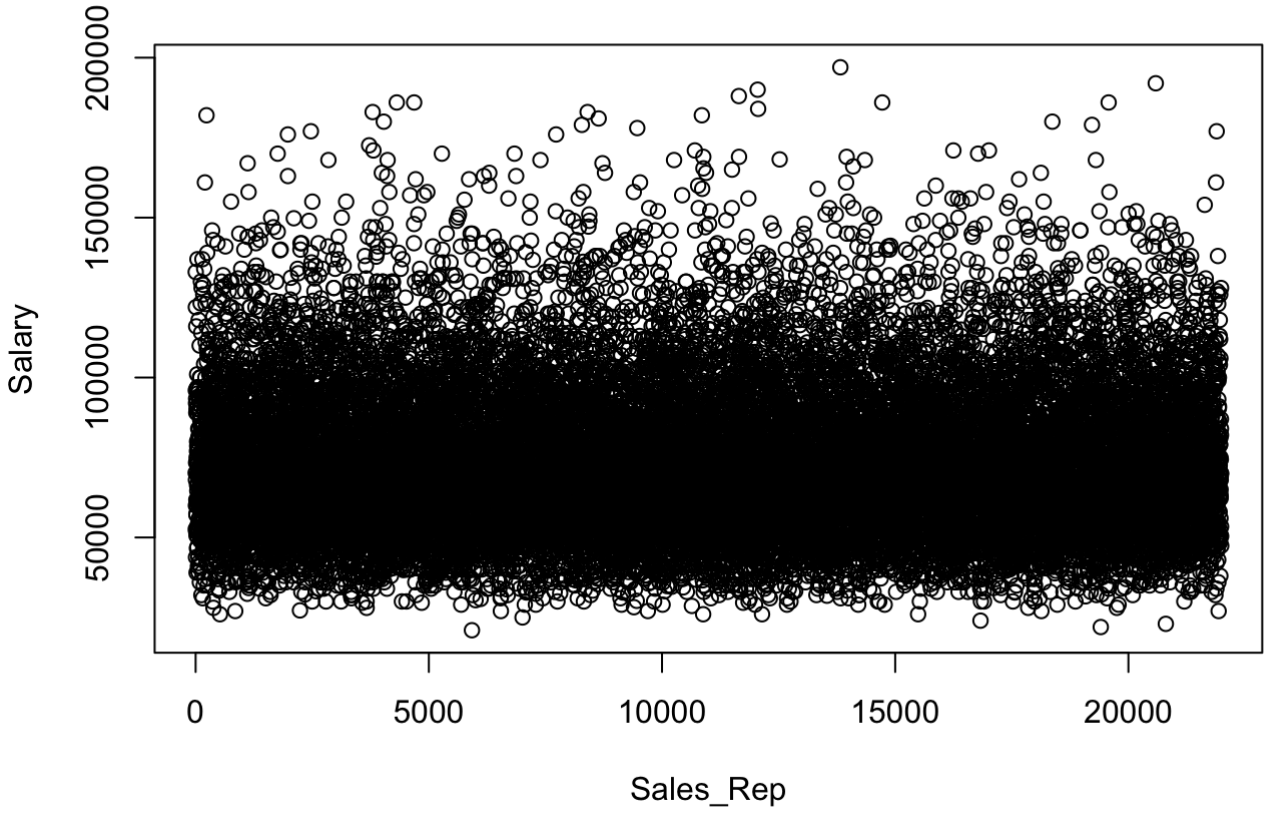
```
plot(Age,certficates)
```




```
plot(Age,Salary)
```




```
plot(Sales_Rep,Salary)
```


Salary
50000
100000
150000
200000
0
5000
10000
15000
20000
Sales_Rep